Smooth Coefficient Variation in Spatial Econometric Models: A Performance Comparison of Spatial STAR and GWR

Seong Do Yun
Graduate Research Assistant
Dept. of Agricultural Economics, Purdue University | yun16@purdue.edu

Michael S. Delgado
Assistant Professor
Dept. of Agricultural Economics, Purdue University | delgado2@purdue.edu

Raymond J. G. M. Florax
Professor
Dept. of Agricultural Economics, Purdue University | rflorax@purdue.edu
Dept. of Spatial Economics, VU University Amsterdam, The Netherlands

Selected Paper prepared for presentation at the VII World Conference of the Spatial Econometrics Association, Washington, D.C., July 10-12, 2013

June 21. 2013

Smooth Coefficient Variation in Spatial Econometric Models: A Performance Comparison of Spatial STAR and GWR

Seong Do Yun *
Dept. of Agricultural Economics
Purdue University

Michale S. Delgado
Dept. of Agricultural Economics
Purdue University

Raymond J. G. M. Florax
Dept. of Agricultural Economics, Purdue University
Dept. of Spatial Economics, VU University Amsterdam, The Netherlands

ABSTRACT

The visual equivalence of spatial heterogeneity and spatial dependence in economic data creates the need to incorporate both types of spatial effects in spatial regression models. Two approaches that incorporate both spatial heterogeneity and dependence are: (i) a spatial version of locally weighted semiparametric regression, termed 'geographically weighted regression' (GWR), and (ii) spatial 'smooth transition autoregressive' (STAR) models. The GWR approach captures spatial heterogeneity through smooth spatial transitions, while the STAR model incorporates spatial heterogeneity via a smooth spatial transition process. While GWR is typically implemented semiparametrically and STAR models have been implemented in both a flexible parametric and semiparametric setting, the relative performance of the two techniques has not yet been systematically compared. This paper aims to fill this void. Specifically, we focus on the fact that estimates from both types of models are fundamentally a function of spatial location defined in terms of coordinates. Locational variation in the estimated coefficients is operationalized through the use of a spatial kernel in GWR and by means of a spatial transition function in spatial STAR models. We perform Monte Carlo simulations to investigate the relative performance of these regression models in terms of bias, efficiency, and predictive capacity in finite samples. We extend these results by considering the relative performance of these models under varying degrees of model misspecification to explore the relative advantages and disadvantages of these techniques when the data generating process is unknown.

Keywords: Geographically Weighted Regression (GWR), Smooth Transition Autoregressive (STAR), spatial heterogeneity, semiparametric.

JEL code: C14, C15, C21.

*Correspondence to: Seongd Do Yun, Dept. of Agricultural Economics, Purdue Universityät West Lafayette, IN 47907-2056. E-mail: yun16@purdue.edu

This paper is prepared for The VII World Conference of the Spatial Econometrics Association, Washington D.C., Jul. 10 - 12. 2013.

1. INTRODUCTION

The visual equivalence of spatial heterogeneity and spatial dependence in economic data creates the need to incorporate both types of spatial effects in spatial regression models. Econometric modeling of spatial dependence requires tailored techniques; hence such models have garnered substantial scrutiny (Pede et al., 2009, p. 2.). Most of models in spatial econometrics (Anselin, 1988, 2003, 2006) are originally designed to capture spatial dependence or spatial autocorrelation. Spatial heterogeneity or non-statationary occurs if the modeled relationships vary systematically over space. This variation can be captured in the coefficients (e.g., spatial regimes or trends), the error variance (i.e., heteroskedasticity), the functional form, or some combination of these (Pede et al., 2009, Anselin, 1988, 2006). Much of the literature is restricted to the specification and estimation of homogeneous parameter spatial autoregressive processes, disregarding the option of systematic spatial variation in the coefficients.

Recently, the alternative of spatially varying coefficients has gained more attraction, and several discrete as well as smooth modeling approaches have been suggested. In particular, a spatial version of locally weighted semiparametric regression, termed 'geographically weighted regression' (GWR), has contributed to popularizing the idea of capturing spatial heterogeneity through smooth spatial transitions (Fotheringham et al., 2002). Spatial 'Smooth Transition AutoRegressive' (STAR) models constitute another option to incorporate spatial heterogeneity, with the added advantage of providing an estimate of the spatial transition process itself (Pede et al., 2009). While GWR is typically implemented semiparametrically and STAR models have been implemented in a flexible parametric setting (Pede et al., 2009).

Though both the Spatial STAR and GWR have basically the same purpose - i.e., capturing smooth spatial parameter heterogeneity - the relative performance of these two approaches has not yet been systematically studied. Further, comparison between these two approaches is not straightforward because one is not necessarily a generalization of the other. Rather, the two models originated under different concepts of spatial variation. The spatial STAR model, for example, models spatial heterogeneity through a well-defined spatial transition process. The GWR approach does not have any specific distributional form of heterogeneity or spatial transition. Indeed, each of these differences are potential advantages of each method: the spatial STAR allows the econometrician to directly estimate the spatial transition process as a means of modeling heterogeneity in parameters, while the unrestricted form of parameter heterogeneity in the GWR framework is potentially more robust to model uncertainty or (mis)specification.

It is in light of these relative advantages and disadvantages of both the STAR and GWR models that we undertake our analysis. Specifically, we seek an understanding of the relative performance of the STAR and GWR spatial regression models in a variety of contexts: bias, efficiency, predictive capacity, and general flexibility or robustness to model misspecification. Our metrics for comparison stem from the fact that estimates from both types of models are fundamentally a function of spatial location defined in terms of coordinates. Locational variation in the estimated coefficients is operationalized through the use of a spatial kernel in GWR and by means of a spatial transition function in spatial STAR models. Discernible spatial patterns in the estimation results can be rigorously compared through measures of predictive performance and mapping.

We explore a variety of Monte Carlo setups to compare the relative performance of these two estimators. We begin with a simplified data generating process (DGP), and investigate the relative performance of these regression models in terms of bias, efficiency, and predictive capacity. This data generating process is specified as a Spatial STAR rather than GWR since the GWR is expected to be robust to various model specifications given its semiparametric form. This setup allows us to compare the relative performance of these models in the absence of model misspecification that may potentially bias the parametric STAR results. Our results show that the spatial STAR model performs well in capturing heterogeneity in parameters, while the GWR model is best suited for prediction.

We expand our simulations to explore robustness of each model to the spatial weighting structure, model specification, and sample size. Preliminary results suggest that while the performance of the STAR model suffers slightly under misspecification of the spatial transition function, the model generally remains quite robust to model misspecification. This result provides confidence in the flexibility of the STAR specification: even under cases of model uncertainty, the flexibility of the STAR model is sufficient to generally capture the smooth spatial transition process. Further, we expect that analysis of the spatial transition function estimate (e.g., its shape, slope, etc.) is of interest to applied researchers in economics. We emphasize, however, that our results underscore the predictive abilities of the GWR approach: under different model specifications or spatial weighting structures, the predictive accuracy of the GWR framework is remarkable.

The rest of this study begins with an analytic comparison of the two models. This section includes brief introduction of each approach, describes generally known criticisms, and derives an analytic relationship between the two models in the case that the smooth transition function in the STAR framework is known. The next two sections describe the Monte Carlo simulation design and present the simulation results. In the final section, conclusions and further research issues are discussed.

2. ANALYTIC MODEL COMPARISON

2.1 Spatial Heterogeneity: Overview

To anchor our discussion of both spatial STAR and GWR models, we begin with a generalized model of spatial heterogeneity

$$\begin{aligned} y_i &= f(x_i, \varepsilon_i) \\ &= f(x_i) + \varepsilon_i \\ &= f(x_i; \beta_i) + \varepsilon_i \end{aligned} \tag{1}$$

in which the index i refers to a spatial unit of observation, and $f(\cdot)$ is a functional relationship which explains the value of the dependent variable y_i in terms of a vector of independent variables x_i, a vector of parameters β_i, and an error term ε_i. For now, we leave ε_i unspecified so as to allow for arbitrary forms of spatial correlation or distribution (or distributional moments). In Equation (1), we impose the common assumptions that the error term ε_i is additively separable from the regression function, and that the regression function is at least partially specified to include a vector of parameters β_i that vary across observation.

It is clear that in Equation (1), heterogeneity is defined as variation in the parameter vector across spatial observation, and that Equation (1) nests several common models. If we assume that there is no variation in parameters across observations, $\beta_i = \beta \ \forall i$, and that $f(\cdot)$ is linear in parameters, then Equation (1) is reduced to the classical linear in parameters setup $y_i = x_i\beta + \varepsilon_i$. However, these restrictions need not be imposed: in the case of the STAR model, we can define $f(\cdot)$ in terms of a transition function, and in the case of the GWR model, we can define $f(\cdot)$ in terms of nonparametric parameter functions that vary across space.

We will make the following assumptions in our review of GWR and spatial STAR. First, assume that both models have spatial observations based upon the same coordinate system. Coordinates can be based on any system of two or three dimension, such as Cartesian, polar, and spherical coordinates. For simplicity, think of $s_i = (u_i, v_i)$ as two dimensional Cartesian coordinates for observation i at (u_i,v_i). Additionally, suppose

$$\begin{cases} i = 1, 2, \ldots, n \\ y \text{ is a scalar dependent variable} \\ x \text{ is a } k\text{-dimension vector of independent variables} \\ \mu \sim i.i.d.N(0,1) \end{cases}$$

is common to both of models.

2.2 Geographically Weighted Regression (GWR)

McMillen (1996) and Brundson et al. (1996) suggest using locally weighted nonparametric (semiparametric) methods to capture spatial variation, relabeled 'geographically weighted regression' by Brundson et al. (1996). The GWR framework is a straightforward generalization of the classic linear regression model (see, e.g., Fotheringham et al. (2002, p. 52)):

$$y_i = \beta_0 + \sum_k \beta_k x_{ik} + \mu_i.$$

GWR extends this traditional regression framework by allowing local rather than global parameters, so that the model is rewritten as:

$$y_i = \beta_0(u_i, v_i) + \sum_k \beta_k(u_i, v_i) x_{ik} + \mu_i \tag{2}$$

in which the parameters β are now functions of the coordinates (u_i, v_i).

Assume the spatial relation that data points closer to each regression point i are weighted more in the local regression than are data points farther away. Then, defining W_i to be an $n \times n$ diagonal matrix of spatial weights, Equation (2) can be written in weighted least squares form:

$$W_i^{1/2} y = W_i^{1/2} X \beta_i + W_i^{1/2} \varepsilon \tag{3}$$

which leads to the weighted least squares estimator

$$\widehat{\beta}_i = (X' W_i X)^{-1} X' W_i y. \tag{4}$$

Following standard nonparametric techniques, the spatial weights W_i is defined as a kernel function that depends on both the spatial coordinates and a bandwidth q. Let $d_{ij} = d(s_i, s_j)$ be the Euclidean norm between observation i and j, and further define q to be a bandwidth identifying the maximum number of neighbors admitted into the search radius in the estimation of the model at a given location. Cho et al. (2010) indicates that the following four kernel weights are legitimate kernel functions which satisfy general kernel function properties of nonparamatric regressors

$$\begin{cases} \text{Triangular or Bartlett kernel: } w_{ij} = 1 - (d_{ij}/q) \\ \text{Epanechnikov kernel: } w_{ij} = 1 - (d_{ij}/q)^2 \\ \text{Bisquare kernel: } w_{ij} = (1 - (d_{ij}/q)^2)^2 \\ \text{Tricube kernel: } w_{ij} = (1 - (d_{ij}/q)^3)^3 \end{cases}$$

while more kernels can be referenced in Fotheringham et al. (2002, pp. 56–59.). Since the bandwidth q is generally unknown, it is generally selected via nonlinear optimization before implementing Equation (4). The cross-validation procedure is the most general method for obtaining the bandwidth q:

$$\underset{q}{\text{ArgMin}} \quad \sum_{i=1}^{n}[y_i - \widehat{y}_i(q)]^2 \tag{5}$$

as this procedure works by choosing the optimal q so that the average squared error of the model is minimum.

Following Cho et al. (2010), the major concerns stated in GWR can be summarized as follows. First, no explicit spatial structure is modeled. GWR uses the spatial kernel weights to generically model the spatial structure. Thus, there is no explicit spatial structure included in the regression, such as a spatial weights matrix in standard spatial econometrics. Second, artificial spatial clustering is adopted rather than an actual attribute of the data generating process. The cluster by spatial kernel is determined by the researcher not by data themselves. Finally, GWR is sensitive to collinearity since in a repeated weighted least squares process there is a higher chance of collinearity problems.

2.3 Family of Spatial STAR

Florax et al. (2009) and Pede et al. (2009) suggests spatial version of Smooth Transition AutoRegressive (STAR) model. In time series econometrics, STAR regressions model the non-linear dynamics which relax the assumption that parameters associated with the data generating processes are fixed through the series, allowing for the endogenous determination of structural discontinuities across time. As implied by its name, the STAR framework allows model parameters to take on different values across regimes, following a potentially smooth transition process (Florax et al. (2009)). Because Spatial STAR can generate smooth varying coefficient, Florax et al. (2009) and Pede et al. (2009) said that Spatial STAR could be an alternative of GWR. Since Spatial STAR can be easily extended to basic spatial econometric models such as Spatial Lag, Spatial Error, and Spatial ARAR, we will start from the simplest form to extensions. The simplest Spatial STAR models can be written as:

$$y = X\beta + X\delta \circ G(\cdot) + \mu \tag{6}$$

where $G(\cdot)$ is transition function and $\circ$ is Hadamard product indicating element by element multiplication. Note that the transition function is bounded between zero and one, the coefficient is bounded by $[\beta_k, \beta_k + \delta_k]$. The next three are the Spatial STAR version of Spatial ARAR, Spatial Lag, and Spatial Error models.

Spatial ARAR STAR Model:

$$y = \rho Wy + X\beta + X\delta \circ G(\cdot) + (I - \lambda W)^{-1}\mu \tag{7}$$

Spatial Lag STAR Model:

$$y = \rho Wy + X\beta + X\delta \circ G(\cdot) + \mu \tag{8}$$

Spatial Error STAR Model:

$$y = X\beta + X\delta \circ G(\cdot) + (I - \lambda W)^{-1}\mu \tag{9}$$

The most popular transition functions are Logistic and Exponential function by convention of time series STAR model. There is, however, no standard way to determine which a transition function is proper. Any function which is smooth, continuous, right cumulative, and bounded between 0 and 1 can be used as transition function $G(\cdot)$. Let $z_i = z(s_i$ be a proper transition variable in the model. From the structural change literature by Holt and Balagtas (2009), three popular functions are shown in the below.

- Logistic Function (First Order): $G(z;\gamma,c) = [1+exp(-\gamma(z-c)/\sigma_z)]^{-1}$
- Exponential Function (Second Order): $G(z;\gamma,c) = 1+exp[-\gamma(z-c)^2/\sigma_z^2]$
- n-th Order Logistic Function: $G(z;\gamma,c_1,c_2,c_3) = \left[1+exp\left(-\gamma\prod_{i=1}^{n}(z-c_i)/\sigma_z\right)\right]^{-1}$

In the Logistic transition function, γ is the speed-of-adjustment parameter that determines how quickly the model shifts from one regime to another, c is the centrality parameter that determines at what point in the sample the structural change is 50% complete, and σ_z is the standard deviation of the normalized trend variable. Dividing γ by σ_z renders the speed-of-adjustment parameter unit free. For the rest two variables, the role of each parameters are not much different. Each of Spatial STAR model from Equation (6) to Equation (9) can be estimated by Nonlinear Least Squares. Especially, Florax et al. (2009) suggest Concentrated Maximum Likelihood Estimation for Equation (7) and LM tests to Equation (9) by the first Taylor approximation for the first stage equation.

There are some realizable concerns about Spatial STAR model. First, Spatial STAR model assumes the fixed functional form including transition function. Thus, there are more chance to be misspeficied in functional form that is generally unknown. Second, estimation is computationally cumbersome. Nonlinear models are generally computationally expensive. If we want to use ML approach, then there are possibility of well known operational burden in huge matrix. The final concern is that the numerical solutions are sensitive to initial values as general in nonlinear function.

Based on concerns in two models, the critics on each model can be summarized in the below.

GWR	Spatial STAR
· No explicit spatial structure	· Fixed functional form
· Artificial spatial clustering	· Specification of G function
· Possibility of extreme coefficient	· Computationally cumbersome
· Sensitive to collinearity	· Sensitive to initial values

3. MONTE CARLO SIMULATION ANALYSIS

As noted before, derivation of explicit comparison standard for both models are cumbersome. The only noticeable common fact is that both models are capturing spatial variation based on spatial locations. Assume that we use a certain coordinate system represented by s. Then, the two coefficient vectors are conceptually expressed as:

- $\alpha_{STAR} = \beta + \delta \circ G(s)$
- $\alpha_{GWR} = \beta(s)$

Since no formal way to compare these two distributions of coefficients has not been suggested, this study is trying to fill this void based upon Monte Carlo Simulation performances.

Two Monte Carlo simulation are designed to investigate model performances with respect to bias, efficiency, and predictive capacity. The first simulations are planed to give illustrative comparison through simple experiments. The second simulations are designed for comparing results of possible combinations of spatial weights matrics, DGP, and estimation function. In both of simulations, the experimental design follows standard practice for Monte Carlo simulations typically used in a spatial econometric context (Anselin and Rey (1991); Anselin and Florax (1995)).

At first, the lattice type geography is generated with (x, y) Euclidean coordinates. Coordinates are assigned by all combinations of integers as $x, y \in [1, n]$. The $n \times n$ synthetic lattice geography is represented by the Figure 1 with n=20, 20×20 regions.

Figure 1 is about here: see the end of the paper

3.1 Simulation I

The first simulation is intended to give an illustrative examples. For this purpose, experimental design is given as simple as possible. Since two models are assuming different schemes of Data Generation Process (DGP), special consideration is required to reduce penalties from DGP. If one model has too strong penalty from data generation from the other, it is harder to compare model performances. Thus, this study focus on the benefit of nonparametric approach. Since GWR is non(semi)parametric approach, it is more flexible to capture spatial variation. With this assumption, DGP of Spatial STAR is applied into Equation (6) as:

$$y_i = 1 + x_i + \frac{1 + x_i}{1 + exp\left(\frac{-5(Wx - \mu_{Wx})}{\sigma_{Wx}}\right)} + \mu_i \tag{10}$$

where $x_i \sim U(0, 1)$ and $\mu_i \sim N(0, 1)$. As we can see, transition function $G(\cdot)$ is defined as Logistic function with $\beta_0 = \beta_1 = \delta_0 = \delta_1 = 1$, and $\gamma = 5$. To avoid another penalty from spatial weights matrix or spatial kernel, spatial kernel for GWR is used as spatial weights matrix in Spatial STAR. Using the tricube kernel with bandwidth $q = 5$, the weights w_{ij} is defined as:

$$w_{ij} = \left(1 - \left(\frac{d_{ij}}{q_{ij}}\right)^3\right)^3 I(d_{ij} < q_{ij}) \tag{11}$$

where $I(\cdot)$ is an index function. From the DGP equation, we assume that x is governs the regime transition, the transition variable $s = Wx$, $c = \mu_{Wx} = 1.0266$, and $\sigma_{Wx} = 0.3464$. For numerical estimation, Gauss-Newton nonlinear least squares are used for Spatial STAR and Equation (4) is applied to GWR. Total 500 replications are implemented. Table I shows descriptive statistics of estimates and failure rates.

Table I is about here: see the end of the paper

In this simulation, we fix a explanatory variable x_i randomly generated from $U(0,1)$ and we assumed that the bandwidth is known as $q = 5$ for all 500 replications. Thus, only the error term μ_i is randomly generated and no cross validation is fulfilled in GWR. To capture divergence or singularity in spatial STAR, starting values are fixed at the true parameter values such as $\beta_0 = \beta_1 = \delta_0 = \delta_1 = 1$, and $\gamma = 5$. The GWR estimates are mean values for each replication and failures are the ratio of the non-full rank matrices of $\sqrt{w_i}x_i$ among 500 replications. As shown in the table, all estimates in spatial STAR seed to be close to the true values while the means of estimates in GWR varies with different ranges. Among 500 replications, we choose two estimates for a well estimated case and a badly estimated case. Figure 2 shows the scatter plot results of those cases.

Figure 2 is about here: see the end of the paper

Using the median of estimates, contour plots are drawn as shown in Figure 3 for β_0, Figure 4 for β_1, and Figure 5 for the predicted values $\widehat{y}$. In case of spatial STAR, β_0 means $\beta_0 + \delta_0 G(\cdot)$ and β_1 means $\beta_1 + \delta_1 G(\cdot)$. In GWR, all estimated and predicted values are the averages of replications.

Figure 3 is about here: see the end of the paper
Figure 4 is about here: see the end of the paper
Figure 5 is about here: see the end of the paper

From the Monte Carlo simulations and illustrative datasets, Spatial STAR is relatively well performed in capturing variations of parameters while GWR is quite better in prediction. GWR shows better performance for getting prediction values. From the scatter plots and mapping result of Figure 13, GWR is very well capturing predicted values which is close to the true values. In case of Spatial STAR, it seems to have upper bound to capture predicted values. All scatter plots of Spatial STAR is upper bounded in some sense. We, however, have no reasonable explanation for this.

Spatial STAR is better performed to get estimates. From the 500 replication of Monte Carlo simulations, the median values of all estimates are very close to the true parameter value. Thus, if the model specification is correct, Spatial STAR could be a better way of getting proper estimates. All scatter plots of Spatial STAR indicates that estimates are smoothly fitted to the actual transition line even in the case of over- or under- estimated dataset. On the other hand, scatter plot of GWR represents that estimates are quite widely distributed. This may because GWR is not assumed any distributional function. Thus, all scatter points in the scatter plots are point-wise marginal changes.

The mapping plots shows that GWR and Spatial STAR have similar spatial variation pattern in estimates and predicted values even though size of values are quite different. This study argue that this may mean two models are using different weighting scheme to capture spatial variation. Since two models are numerically different methods, the difference of size may come from usage different weighting schemes. Thus, comparison of two models could be better interpreted with spatial variation pattern rather than magnitude of estimates or predicted values.

3.2 Simulation 2

In this section, simulations are designed to capture estimation performances under different combinations of spatial weights matrices, Data Generating Processes, and estimation functions. We choose the Means Squared Error (MSE) and the Mean Absolute Deviation Error (MADE) for each coefficient function over 2,000 replications. MSE and MADE are defined as:

$$MSE = \frac{1}{m}\sum_{m=1}^{M}\left(\theta_i(s_i) - \widehat{\theta}_i(s_i)\right)^2 \tag{12}$$

$$MADE = \frac{1}{m}\sum_{m=1}^{M}|\theta_i(s_i) - \widehat{\theta}_i(s_i)| \tag{13}$$

where m is the number of replications, θ_i is the true value at the location s_i, and $\widehat{\theta}_i$ is the estimate. These two measures are used for estimates and predictive values. To get MSE and MADE of estimates in GWR, we focus on the relationship between spatial STAR DGP and functional form of GWR. Let's assume that we are using Equation 6 as DGP and Equation 3 as weighted least squares form of GWR with the tricube weights of Equation 11. After rearranging spatial STAR, we can get

$$\begin{cases} y_i &= (\beta_0 + \delta_0 \circ G) + (\beta_1 + \delta_1 \circ G)x_i \\ \sqrt{W_i}y &= \sqrt{W_i}X\beta \end{cases} \tag{14}$$

Consider that $\sqrt{W_i}$ is weighting and selecting data in GWR. Then, the similar role is done by $G(\dot{)}$ which is function of W_i in spatial STAR. Thus, we can simply set up the relation that $(\beta_0 + \delta_0 \circ G)$ in spatial STAR is comparable to intercept estimates in GWR. And $(\beta_1 + \delta_1 \circ G)$ is comparable to $\beta(s_i)$ in GWR. Thus, the calculated two values in spatial STAR DGP are treated as true value of estimates in GWR. In all simulation tables in this section, MSE and MADE of estimates in GWR are calculated with this way.

At the simulations in this section, x_i is randomly generated from $U(0,1)$ for each replication and the bandwidth q is chosen from the cross validation of Equation 5. For spatial STAR simulations, we adopt multi-initial value techniques with the derivative-free optimization algorithms by quadratic approximation (Powell, 2009). This can help to reduce divergence and singular gradient problem in spatial STAR. Ten vectors are randomly chosen as multi-initial values as $\beta_0, \beta_1, \delta_0, \delta_1 \sim U(0,2)$, $\gamma \sim U(4,6)$, and $c \sim U(\mu_{Wx} - 0.5, \mu_{Wx} + 0.5)$. Total 2,000 replications are applied to all of simulations.

Table II shows simulation results from different sample sizes with $n = 25, 100, 400, 900$, and 1600. The logistic DGP in Equation 10 and logistic estimation function of $G(\cdot)$ are used for 2,000 replications.

Table II is about here: see the end of the paper

In the spatial STAR, MSE and MADE of estimates show that the bias becomes smaller as the sample size increases while the predictive values are keep the similar level of MSE and MADE. Comparing to the spatial STAR, GWR shows lower MSE and MADE for the averages of predictive values.

Table III shows 2,000 replications of different Data Generating Functions with the Logistic $G(\cdot)$ estimation function in the spatial STAR. For comparison, two additional DGPs, Exponential and Second order logistic, are additionally generated. We assumed Exponential DGP as:

$$y_i = 1 + x_i + (1 + x_i)exp\left[-\left(\frac{\gamma(Wx - \mu_{Wx})^2}{\sigma^2_{Wx}}\right)\right] + \mu_i \tag{15}$$

The Second Order Logistic DGP is defined as:

$$y_i = 1 + x_i + \frac{1 + x_i}{1 + exp\left(\frac{-5(Wx-0.4)(Wx-0.7)}{\sigma^2_{Wx}}\right)} + \mu_i \tag{16}$$

Table III is about here: see the end of the paper

As shown in Table III, "Exponential DGP + Exponential $G(\cdot)$ Estimation" shows the best performance in spatial STAR results.

To depict the opposite situation, Table IV implements 2,000 replications of the Logistic DGP with the Exponential $G(\cdot)$ estimation function in the spatial STAR.

Table IV is about here: see the end of the paper

Table IV shows that "Logistic DGP + Logistic $G(\cdot)$ Estimation" is better than the other combination. The results from Table III and Table IV are graphically summarized to the Figure 6.

Figure 6 is about here: see the end of the paper

In the previous simulation, we fix the spatial weights (or kernel) as Tricube, Equation 11. The final simulations adopt a spatial weights concept which is general in spatial STAR. The queen spatial weights matrix is defined as:

$$w_{ij} = \begin{cases} 1 & \text{if } i \text{ and } j \text{ is connected} \\ 0 & \text{otherwise} \end{cases} \tag{17}$$

The weights matrix in DGP is row standardized. Table V shows 2,000 replications of the spatial STAR weights and Logistic DGP with the Exponential $G(\cdot)$ estimation function in the spatial STAR.

Table V is about here: see the end of the paper

It is interesting that failure rate in the spatial STAR is increased while it is zero in GWR. As noted, we adopt multi-initial value approach in spatial STAR to reduce failure rate of estimation which comes from divergence or singular gradient. This increase of failure rate means that ten initial values in some replications is not enough to get convergent values. In case of GWR, multicollinearity is eliminated from adopting spatial STAR weights for this specific simulations. As we expected, "Logistic DGP + Logistic $G(\cdot)$ Estimation" in spatial STAR shows better performance than "Logistic DGP + Exponential $G(\cdot)$ Estimation".

4. CONCLUSIONS

Much of the spatial econometric literature incorporating spatially varying coefficients is restricted to the specification and estimation of homogeneous parameter spatial autoregressive processes, disregarding the option of systematic spatial variation in the coefficients. Recently, the alternative of spatially varying coefficients has gained more traction, and two alternatives, GWR and Spatial STAR seem to be representative among those. While GWR is typically implemented semiparametrically and STAR models have been implemented in both a flexible parametric and semiparametric setting, the relative performance of the two techniques has not yet been systematically compared. This paper aims to fill this void by comparing estimation performances. Because of difference in estimation concepts, defining explicit comparison standard between two models are not easy. This study specifically focus on the fact that estimates from both types of models are fundamentally a function of spatial location defined in terms of coordinates. Locational variation in the estimated coefficients is operationalized through the use of a spatial kernel in GWR and by means of a spatial transition function in spatial STAR models. Discernible spatial patterns in the estimation results can be fruitfully compared through mapping. Starting from a simplified data generating process (DGP), we perform Monte Carlo simulations to investigate the relative performance of these regression models in terms of bias, efficiency, and predictive capacity in finite samples. And the results are summarized by the following four.

First, GWR shows better performance for getting prediction values. From the mapping results and 2,000 Monte Carlo simulations, predicted values of GWR simulates very close to the true values. Since GWR using more data points to produce a predicted value, this results seems to be reasonable. Thus, GWR could be appropriate to get accurate predicted value.

Second, Spatial STAR is better performed to get estimates. From the replications of Monte Carlo simulations, the median values of all estimates are very close to the true parameter value. Thus, if the model specification is correct, Spatial STAR could be a better way of getting proper estimates. However, it shows 12% in the simulation 1 and 3.85% in the Table V of failure rates from convergence or non-full rank gradient problem. This means that setting proper initial values or considering several alternative of estimation methods are often necessary.

Third, GWR and Spatial STAR are using different weighting scheme to capture spatial variation. Even though each model has better performance in estimates or predicted values, spatial variation patterns on the map seems to be similar. Since two models are numerically different methods, the difference of size may come from usage different weighting schemes. Thus, comparison of two models could be better interpreted with spatial variation pattern rather than magnitude of estimates or predicted values.

At last, choosing proper estimation function $G(\cdot)$ in spatial STAR seems important. Table III and Table IV shows that only correct specification cases supports the smallest bias and the best prediction capacity. In the most of empirical study, real GDP is often unknown and researchers need to choose estimation function by certain standards. Thus, nonparametric approaches to select the correct $G(\cdot)$ function or nonparametric spatial econometric approach could be considered in practical manner.

In conclusion, Spatial STAR is relatively well performed in capturing variations of parameters while GWR is quite better in prediction. Lastly, we can not avoid to say some limitations of this paper. Even though this study seem to reach reasonable simulation results, no rigorous spatial comparison of spatial distribution between GWR and spatial STAR is used. Syrjala's test which is an approximate randomization test for comparing two spatial distribution Syrjala (1996) could be implemented in the later version of this paper.

REFERENCES

Anselin L. 1988. *Spatial Econometrics: Methods and Models*. Dordrecht: Kluwer Academic Publishers.

Anselin L. 2003. Spatial externalities, spatial multipliers, and spatial econometrics. *International Regional Science Review* **26**: 153–166.

Anselin L. 2006. Spatial econometrics. In Mills TC, Patterson K (eds.) *Palgrave Handbook of Econometrics*, Vol. 1 Econometric Theory. Basingstoke: Palgrave Macmillan, 901–969.

Anselin L, Florax RJGM. 1995. Small sample properties of tests for spatial dependence in regression models: Some further results. In Anselin L, Florax RJGM (eds.) *New Directions in Spatial Econometrics*. Berlin: Springer, 21–74.

Anselin L, Rey S. 1991. Properties of tests for spatial dependence in linear regression models. *Geographical Analysis* **23**: 112–131.

Brundson CA, Fotheringham AS, Charlton ME. 1996. Geographically weighted regression: A method for exploring spatial nonstationarity. *Geographical Analysis* **28**: 281–298.

Cho SH, Lambert D, Chen Z. 2010. Geographically weighted regression bandwidth selection and spatial autocorrelation: an empirical example using chinese agriculture data. *Applied Economics Letters* **17**: 767–772.

Florax RJGM, Pede VO, Lambert DM, Holt MT. 2009. Spatial econometric star models: Lm tests and monte carlo simulations. Presented at the Annual Meetings of the 56th North American Regional Science Association, 2009 NARSC: San Francisco, California. Nov. 19-21, 2009.

Fotheringham AS, Brundson CA, Charlton ME. 2002. *Geographically Weighted Regression: the Analysis of Spatially Varying Relationships*. West Sussex: John Wiley & Sons.

Holt MT, Balagtas JV. 2009. Estimating structural change with smooth transition regressions: An application to meat demand. *American Journal of Agricultural Economics* **91**: 1424–1431.

McMillen DP. 1996. One hundred fifty years of land values in chicago: A nonparametric approach. *Journal of Urban Economics* **40**: 100–124.

Mei CL, Wang N, Zhang WX. 2006. Testing the importance of the explanatory variables in a mixed geographically weighted regression model. *Environment and Planning A* **38**: 587–598.

Pede VO, Florax RJGM, Holt MT. 2009. Spatial econometric star model with an appication to u.s. county economic growth. URL http://ageconsearch.umn.edu/bitstream/48117/2/09-03.pdf. Working Paper 09-03, Dept. of Agricultural Economics, Purdue University.

Powell MJD. 2009. The bobyqa algorithm for bound constrained optimization without derivatives. Report DAMTP 2009/NA06, Centre for Mathematical Sciences, University of Cambridge, UK. URL http://www.damtp.cam.ac.uk/user/na/NA_papers/NA2009_06.pdf

Syrjala SE. 1996. A statistical test for a difference between the spatial distributions of two populations. *Ecology* **77**: 75–80.

Table I: Monte Carlo Simulation Results (500 Replications)

	GWR		Spatial STAR					
	β_0	β_1	β_0	β_1	δ_0	δ_1	γ	c
Min.	0.4607	2.3780	0.5081	-2.5075	-0.3135	-2.3066	1.3740	0.6332
1st Q.	0.6988	2.9340	0.8846	0.5090	0.7241	0.3617	3.9950	1.0044
Median	0.7877	3.0690	1.0032	1.0024	1.0049	0.9923	5.0910	1.0326
Mean	0.7866	3.0820	0.9995	0.9461	1.0319	1.0318	6.4040	1.0308
3rd Q.	0.8714	3.236	1.1163	1.4130	1.2970	1.6695	6.7450	1.0604
Max.	1.2789	3.670	1.5451	2.6222	3.9422	5.8687	93.3720	1.4416
Failure	18.23%		12.00%					

1. 1st Q. and 3rd Q. mean 25% and 75% Quartiles.
2. Failure means ratio of failures to get estimates among 500 replications.

Table II: Monte Carlo Simulation Results: Different Sample Size with Logistic DGP (2,000 Replications)

	GWR					Spatial STAR								
n=25	q	β_0	β_1	$\widehat{y}$ MSE	$\widehat{y}$ MADE	β_0	β_1	δ_0	δ_1	γ	c	True c	$\widehat{y}$ MSE	$\widehat{y}$ MADE
Min.	5.0000	-1.8751	-1.0562	0.0413	0.1439	-9.3689	-20.3178	-7.1974	-15.6315	0.4794	0.0562	0.6728	0.1518	0.2967
1st Q.	11.0000	0.3995	2.6404	0.4675	0.5360	0.5984	-0.1100	0.0444	-0.5895	4.0127	0.9336	0.9361	0.6756	0.6579
Median	17.0000	0.7105	3.2074	0.6605	0.6467	0.9638	1.1113	0.9805	1.0868	5.3113	1.0603	1.0194	0.8634	0.7462
Mean	15.2753	0.7073	3.1949	0.6853	0.6390	0.9364	0.8318	1.2341	1.2027	5.2929	1.0616	1.0212	0.9641	0.7647
3rd Q.	20.0000	1.0212	3.7591	0.8835	0.7515	1.3597	2.2762	2.1056	2.7397	6.4056	1.1889	1.0986	1.0987	0.8459
Max.	20.0000	3.1002	7.7530	2.0468	1.1759	4.0037	7.4654	22.6711	23.9205	15.0873	2.1976	1.4246	23.7756	2.4171
MSE		1.8205	6.5750			0.6370	5.3669	4.5563	10.4698	3.9791	0.0341			
MADE		1.0124	2.0123			0.4987	1.5967	1.4124	2.2808	1.5422	0.1375			
Failure	3.20%					0.00%								
	GWR					Spatial STAR								
n=100	q	β_0	β_1	$\widehat{y}$ MSE	$\widehat{y}$ MADE	β_0	β_1	δ_0	δ_1	γ	c	True c	$\widehat{y}$ MSE	$\widehat{y}$ MADE
Min.	5.0000	-0.9326	-1.9775	0.0882	0.2344	-2.1343	-16.1834	-2.9693	-7.3551	0.4706	0.3469	0.8200	0.5529	0.5982
1st Q.	11.0000	0.5560	2.9418	0.5095	0.5634	0.7888	0.0770	0.4536	-0.0500	3.8269	0.9373	0.9653	0.8656	0.7406
Median	17.0000	0.7240	3.2197	0.6592	0.6415	0.9933	0.9218	1.0697	1.0584	5.2294	1.0151	1.0073	0.9599	0.7830
Mean	14.9739	0.7188	3.2191	0.6277	0.6128	0.9779	0.7695	1.1789	1.1570	5.0296	0.7695	1.0076	0.9713	0.7857
3rd Q.	20.0000	0.8836	3.4940	0.7782	0.7012	1.1758	1.6679	1.7286	2.1776	6.1873	1.6679	1.0490	1.0613	0.8280
Max.	20.0000	2.5210	5.0879	1.3252	0.9235	2.5767	5.6760	14.1449	19.2166	12.6210	5.6760	1.1935	4.1502	1.3463
MSE		2.1025	7.8490			0.1150	2.4978	1.3448	4.4307	2.9634	0.0198			
MADE		0.9980	1.9943			0.2450	1.0575	0.8111	1.4679	1.3756	0.0947			
Failure	4.15%					0.00%								
	GWR					Spatial STAR								
n=400	q	β_0	β_1	$\widehat{y}$ MSE	$\widehat{y}$ MADE	β_0	β_1	δ_0	δ_1	γ	c	True c	$\widehat{y}$ MSE	$\widehat{y}$ MADE
Min.	5.0000	0.3265	2.3570	0.1322	0.2695	-0.1097	-4.7861	-1.0158	-2.8123	0.8712	0.3554	0.9225	0.7823	0.6975
1st Q.	8.0000	0.6499	3.0551	0.2655	0.3873	0.9008	0.5096	0.6927	0.5069	3.9709	0.9760	0.9980	0.9420	0.7737
Median	11.0000	0.7317	3.1938	0.5591	0.5871	0.9973	0.9243	1.0106	1.0875	5.1490	1.0155	1.0186	0.9849	0.7937
Mean	10.3409	0.7309	3.1935	0.4886	0.5329	0.9972	0.8777	1.0191	1.1212	5.0594	1.0148	1.0184	0.9895	0.7942
3rd Q.	12.0000	0.8102	3.3379	0.6193	0.6192	1.0941	1.2955	1.3192	1.6770	6.1109	1.0541	1.0392	1.0376	0.8140
Max.	20.0000	1.2079	3.9901	0.9018	0.7468	1.5587	3.6325	2.9560	8.3704	15.5568	1.6787	1.1156	1.2748	0.9079
MSE		1.9673	6.8674			0.0225	0.5145	0.2428	0.9355	2.0730	0.0053			
MADE		1.0520	2.0720			0.1170	0.5161	0.3853	0.7233	1.1728	0.0439			
Failures	3.20%					0.00%								

Continued on next page

Table II – *Continued from previous page*

	GWR					Spatial STAR								
n=900	q	β_0	β_1	$\widehat{y}$ MSE	$\widehat{y}$ MADE	β_0	β_1	δ_0	δ_1	γ	c	True c	$\widehat{y}$ MSE	$\widehat{y}$ MADE
Min.						0.3246	-1.1431	0.0473	-1.7168	2.2485	0.4366	0.9645	0.8350	0.7284
1st Q.						0.9362	0.7155	0.7944	0.6507	4.3336	1.0004	1.0115	0.9612	0.7828
Median						0.9942	0.9926	1.0058	1.0075	5.1212	1.0239	1.0242	0.9939	0.7967
Mean						0.9965	0.9701	1.0180	1.0195	5.1692	1.0239	1.0247	0.9952	0.7962
3rd Q.						1.0604	1.2449	1.2286	1.3795	6.0258	1.0478	1.0382	1.0279	0.8092
Max.						1.3234	3.3212	2.0320	3.6582	12.2106	1.5888	1.0921	1.1620	0.8663
MSE						0.0089	0.1706	0.1000	0.3253	1.3173	0.0013			
MADE						0.0741	0.3185	0.2519	0.4437	0.9361	0.0244			
Failure						0.00%								
	GWR					Spatial STAR								
n=1600	q	β_0	β_1	$\widehat{y}$ MSE	$\widehat{y}$ MADE	β_0	β_1	δ_0	δ_1	γ	c	True c	$\widehat{y}$ MSE	$\widehat{y}$ MADE
Min.						0.6632	-0.3644	0.1321	-0.3448	2.6615	0.9495	0.9758	0.8733	0.7421
1st Q.						0.9521	0.7924	0.8340	0.7384	4.4507	1.0113	1.0178	0.9727	0.7867
Median						0.9992	0.9949	0.9976	1.0084	5.0608	1.0271	1.0283	0.9965	0.7966
Mean						0.9983	0.9858	1.0009	1.0188	5.1634	1.0277	1.0281	0.9969	0.7966
3rd Q.						0.9983	1.1919	1.1610	1.3049	5.8284	1.0445	1.0379	1.0210	0.8069
Max.						1.2264	3.3258	1.9141	2.7289	8.2669	1.6952	1.0756	1.1129	0.8519
MSE						0.0048	0.0887	0.0580	0.1775	0.9326	0.0007			
MADE						0.0549	0.2352	0.1925	0.3364	0.7813	0.0166			
Failure						0.00%								

1. 1st Q. and 3rd Q. mean 25% and 75% Quartiles.
2. Failure means ratio of failures to get estimates among 2,000 replications.
3. $\widehat{y}$MSE and $\widehat{y}$MADE are calculated at each replication.
4. MSE and MADE in GWR are the average of estimates from 2,000 replications.
5. Two cases of n=900 and n=1600 in GWR will be updated at the later version of this paper.

Table III: Monte Carlo Simulation Results: Exponential DGP with the Logistic Estimation Function (2,000 Replications)

	GWR					Spatial STAR (Exponential DGP + Exponential $G(\cdot)$ Estimation)								
n=400	q	β_0	β_1	$\widehat{y}$ MSE	$\widehat{y}$ MADE	β_0	β_1	δ_0	δ_1	γ	c	True c	$\widehat{y}$ MSE	$\widehat{y}$ MADE
Min.	5.0000	1.2155	1.0142	0.1584	0.2926	-0.3233	-1.5796	-0.8709	-1.6099	0.7031	0.7640	0.9225	0.7823	0.6976
1st Q.	8.0000	1.5180	1.6991	0.3241	0.4263	0.7541	0.5302	0.7143	0.5063	4.1802	0.9946	0.9980	0.9419	0.7739
Median	11.0000	1.6054	1.8434	0.6243	0.6212	1.0180	0.9955	0.9930	1.0074	4.9431	1.0193	1.0186	0.9843	0.7933
Mean	10.5894	1.6015	1.8461	0.5558	0.5698	1.0086	1.0059	0.9956	0.9958	4.9343	1.0191	1.0184	0.9892	0.7940
3rd Q.	12.0000	1.6874	1.9919	0.6897	0.6534	1.2597	1.4696	1.2660	1.5146	5.7336	1.0436	1.0392	1.0335	0.8134
Max.	20.0000	2.0251	2.5915	0.9497	0.7707	2.6238	3.4526	2.3713	3.4657	8.1910	1.4195	1.1156	1.6204	1.0324
MSE		1.4599	4.5952			0.1451	0.5085	0.1688	0.5741	1.1471	0.0006			
MADE		0.8559	1.5469			0.3020	0.5637	0.3259	0.6000	0.8796	0.0149			
Failure	2.70%					0.00%								
						Spatial STAR (Exponential DGP + Logistic $G(\cdot)$ Estimation)								
n=400						β_0	β_1	δ_0	δ_1	γ	c		$\widehat{y}$ MSE	$\widehat{y}$ MADE
Min.						-0.2958	-4.8475	-3.8033	-2.2124	2.2328	0.3651		0.8638	0.7219
1st Q.						2.0609	-0.3876	-0.8837	1.7916	5.7601	1.1849		1.0695	0.8246
Median						2.1523	-0.1312	-0.3813	2.3776	6.3900	1.2303		1.1250	0.8476
Mean						2.1522	-0.0935	-0.4364	2.4198	6.4306	1.2134		1.1451	0.8536
3rd Q.						2.2465	0.1311	0.0733	3.0454	7.0296	1.2653		1.1909	0.8732
Max.						2.7070	2.9260	2.8952	7.3332	15.2956	1.6515		2.6272	1.3184
Failure						0.00%								
	GWR					Spatial STAR (2n Order Logistic DGP + Logistic $G(\cdot)$ Estimation)								
n=400	q	β_0	β_1	$\widehat{y}$ MSE	$\widehat{y}$ MADE	β_0	β_1	δ_0	δ_1	γ	c		$\widehat{y}$ MSE	$\widehat{y}$ MADE
Min.	5.0000	0.1286	2.5533	0.1852	0.3240	-1.4493	-10.5857	-2.9430	-7.5445	0.9746	0.5643		0.9539	0.7601
1st Q.	9.0000	0.6565	3.4389	0.6679	0.6329	-0.0746	-2.8145	1.5643	1.9962	5.1379	0.7532		1.1900	0.8668
Median	11.0000	0.7699	3.6309	0.7844	0.6896	0.1878	-1.5623	1.8640	3.6403	6.0706	0.7710		1.3003	0.9052
Mean	11.1638	0.7709	3.6279	0.7237	0.6463	0.1637	-0.9998	1.8855	3.0521	6.2634	0.8719		1.3474	0.9156
3rd Q.	13.0000	0.8906	3.8163	0.8810	0.7322	0.4231	0.0567	2.1557	4.8923	7.0641	0.7929		1.4469	0.9539
Max.	20.0000	1.3479	4.6053	1.2484	0.8767	1.1454	5.2789	11.6307	20.3600	19.3339	2.0225		11.4826	2.6683
MSE		2.3823	11.2804											
MADE		1.2005	2.5490											
Failure	2.65%					0.00%								

1. 1st Q. and 3rd Q. mean 25% and 75% Quartiles.
2. Failure means ratio of failures to get estimates among 2,000 replications.
3. $\widehat{y}$MSE and $\widehat{y}$MADE are calculated at each replication.
4. MSE and MADE in GWR are the average of estimates from 2,000 replications.

Table IV: Monte Carlo Simulation Results:Logistic DGP with the Exponential Estimation Function (2,000 Replications)

	GWR					Spatial STAR (Logistic DGP + Logistic $G(\cdot)$ Estimation)								
n=400	q	β_0	β_1	$\widehat{y}$ MSE	$\widehat{y}$ MADE	β_0	β_1	δ_0	δ_1	γ	c	True c	$\widehat{y}$ MSE	$\widehat{y}$ MADE
Min.	5.0000	0.3265	2.3570	0.1322	0.2695	-0.1097	-4.7861	-1.0158	-2.8123	0.8712	0.3554	0.9225	0.7823	0.6975
1st Q.	8.0000	0.6499	3.0551	0.2655	0.3873	0.9008	0.5096	0.6927	0.5069	3.9709	0.9760	0.9980	0.9420	0.7737
Median	11.0000	0.7317	3.1938	0.5591	0.5871	0.9973	0.9243	1.0106	1.0875	5.1490	1.0155	1.0186	0.9849	0.7937
Mean	10.3409	0.7309	3.1935	0.4886	0.5329	0.9972	0.8777	1.0191	1.1212	5.0594	1.0148	1.0184	0.9895	0.7942
3rd Q.	12.0000	0.8102	3.3379	0.6193	0.6192	1.0941	1.2955	1.3192	1.6770	6.1109	1.0541	1.0392	1.0376	0.8140
Max.	20.0000	1.2079	3.9901	0.9018	0.7468	1.5587	3.6325	2.9560	8.3704	15.5568	1.6787	1.1156	1.2748	0.9079
MSE		1.9673	6.8674			0.0225	0.5145	0.2428	0.9355	2.0730	0.0053			
MADE		1.0520	2.0720			0.1170	0.5161	0.3853	0.7233	1.1728	0.0439			
Failures	3.20%					0.00%								
						Spatial STAR (Logistic DGP + Exponential $G(\cdot)$ Estimation)								
n=400						β_0	β_1	δ_0	δ_1	γ	c	True c	$\widehat{y}$ MSE	$\widehat{y}$ MADE
min						-0.4294	-2.3222	-3.7404	-1.6345	0.0680	-0.2186		0.7948	0.7028
1st Q.						0.8916	0.1832	-0.3792	1.5604	0.7994	0.6481		0.9625	0.7830
Median						1.0499	0.6774	0.0984	2.2257	1.7836	0.7535		1.0123	0.8038
Mean						1.0917	0.6885	0.2062	2.1832	2.7047	0.7670		1.0218	0.8066
3rd Q.						1.2443	1.1939	0.8212	2.8683	4.2567	0.8563		1.0764	0.8285
Max						4.3535	4.0075	3.3709	5.4873	18.0306	1.7655		2.6995	1.3805
Failure						0.00%								

1. 1st Q. and 3rd Q. mean 25% and 75% Quartiles.
2. Failure means ratio of failures to get estimates among 2,000 replications.
3. $\widehat{y}$MSE and $\widehat{y}$MADE are calculated at each replication.
4. MSE and MADE in GWR are the average of estimates from 2,000 replications.
5. The upper panel results comes from Table II.

Table V: Monte Carlo Simulation Results: Spatial STAR Weights (2,000 Replications)

	GWR				Spatial STAR (Logistic DGP + Logistic $G(\cdot)$ Estimation)								
n=400	β_0	β_1	$\widehat{y}$ MSE	$\widehat{y}$ MADE	β_0	β_1	δ_0	δ_1	γ	c	True c	$\widehat{y}$ MSE	$\widehat{y}$ MADE
Min.	0.97187	-0.6264	1.26779	0.88235	0.1880	-0.0292	-1.7327	-2.1225	-114.3970	0.4414	0.4529	0.7826	0.6992
1st Q.	1.56991	1.01299	1.63317	1.0089	0.9012	0.8541	0.6965	0.4662	3.2460	0.8541	0.4900	0.9379	0.7722
Median	1.64878	1.16357	1.7517	1.04109	1.0351	1.0768	0.9570	0.8880	4.7105	0.4998	0.5002	0.9815	0.7916
Mean	1.64846	1.16663	2.50939	1.04268	1.1554	1.1664	0.6849	0.6681	7.0318	0.4999	0.4999	0.9839	0.7919
3rd Q.	1.73431	1.3162	1.87799	1.07408	1.2010	1.3301	1.1510	1.2134	6.2747	0.5116	0.5103	1.0301	0.8117
Max.	2.15726	2.44167	1326.65	2.78217	2.5436	2.7549	2.0813	3.0912	5686.5569	0.5530	0.5477	1.2424	0.8895
MSE	1.3614	5.0545			0.1942	0.2536	0.7407	0.8701	16848.3183	0.0001			
MADE	0.8249	1.4703			0.2786	0.3586	0.5054	0.6267	6.5979	0.0071			
Failure	0.00%				3.85%								
					Spatial STAR (Logistic DGP + Exponential $G(\cdot)$ Estimation)								
n=400					β_0	β_1	δ_0	δ_1	γ	c		$\widehat{y}$ MSE	$\widehat{y}$ MADE
Min.					-339.6400	-323.8574	-197.9352	-649.6717	0.0553	0.1644		0.9962	0.7965
1st Q.					0.7631	0.6960	0.6665	0.1880	0.3151	0.3407		1.2466	0.8928
Median					0.9331	0.9846	1.1463	1.0672	0.4120	0.3640		1.3192	0.9195
Mean					0.8881	1.7302	0.9518	0.1125	6.4869	0.4210		36.5234	1.1831
3rd Q.					1.2122	1.4900	1.4119	1.5324	0.5290	0.4011		1.4579	0.9758
Max.					199.2710	651.2428	341.0528	325.6586	4687.3178	0.8282		34214.9562	157.0055
Failure					29.60%								

1. 1st Q. and 3rd Q. mean 25% and 75% Quartiles.
2. Failure means ratio of failures to get estimates among 2,000 replications.
3. $\widehat{y}$MSE and $\widehat{y}$MADE are calculated at each replication.
4. MSE and MADE in GWR are the average of estimates from 2,000 replications.

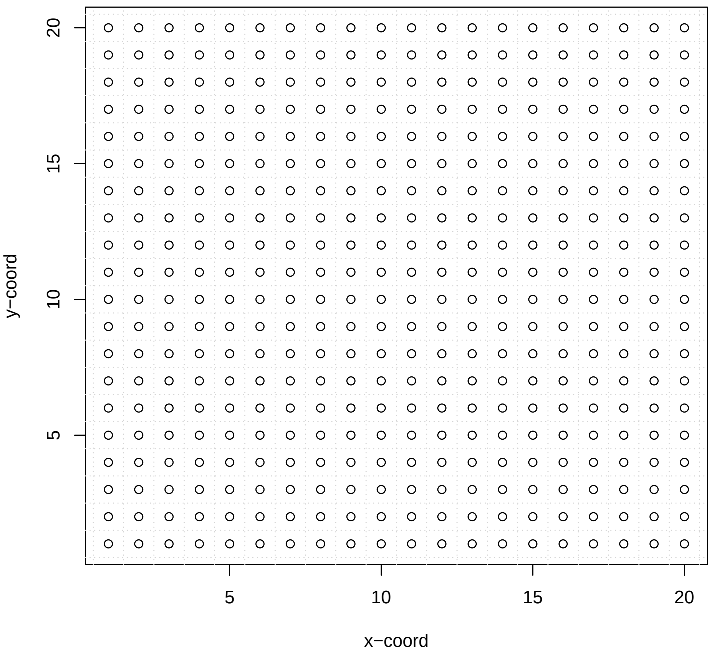

Figure 1: 20×20 Lattice Type Synthetic Geography

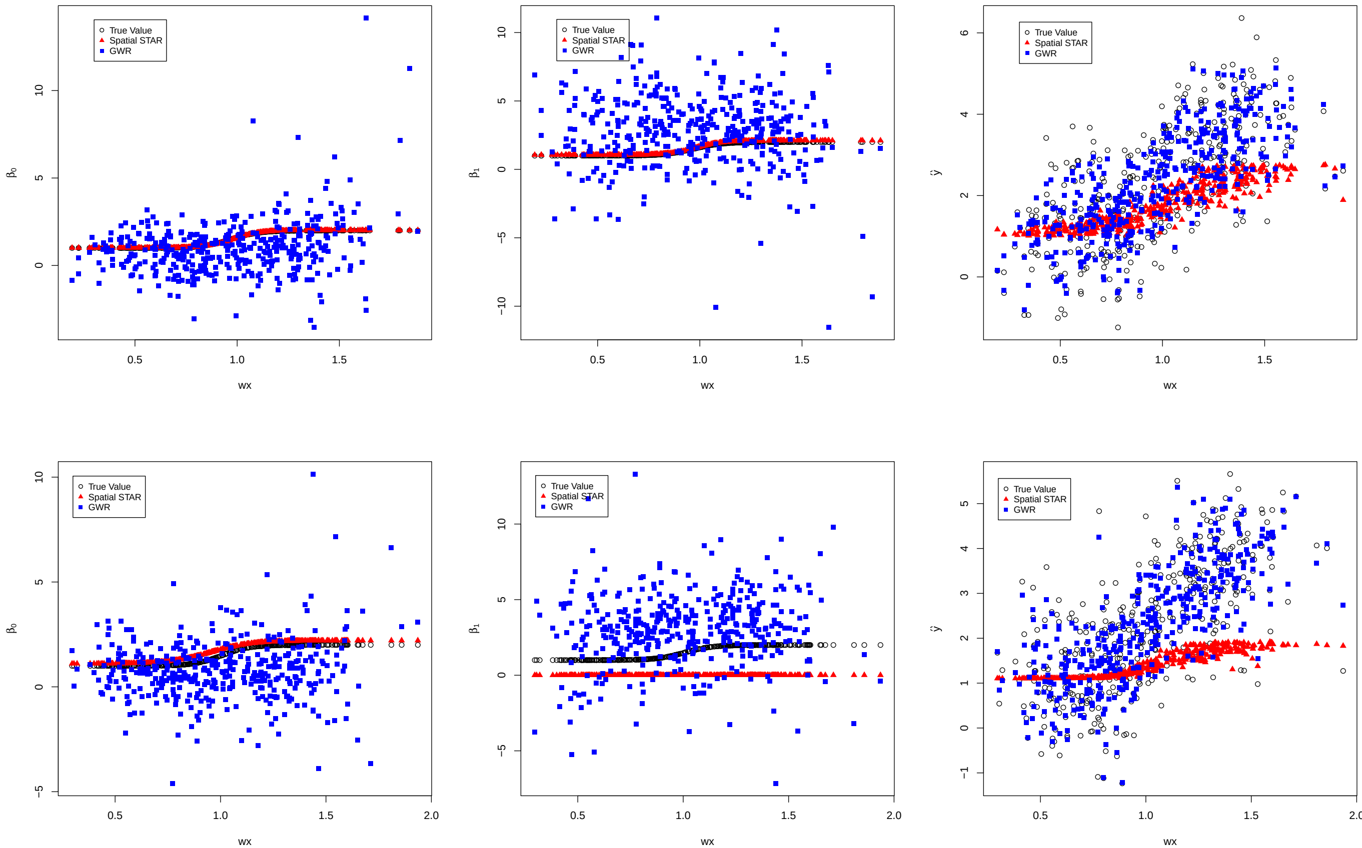

Figure 2: Scatter plots for two selected simulation results: (Upper Panel) Well Estimated Case: $\beta_0 = 1.0087$, $\beta_1 = 1.0539$, $\delta_0 = 1.2462$, $\delta_1 = 0.5614$, $\gamma = 5.3513$, and $\mu_{Wx} = 0.9805$ (Lower Panel) Badly Estimated Case: $\beta_0 = 1.1055$, $\beta_1 = 0.0040$, $\delta_0 = 0.3010$, $\delta_1 = 2.7113$, $\gamma = 3.4181$, and $\mu_{Wx} = 0.9571$

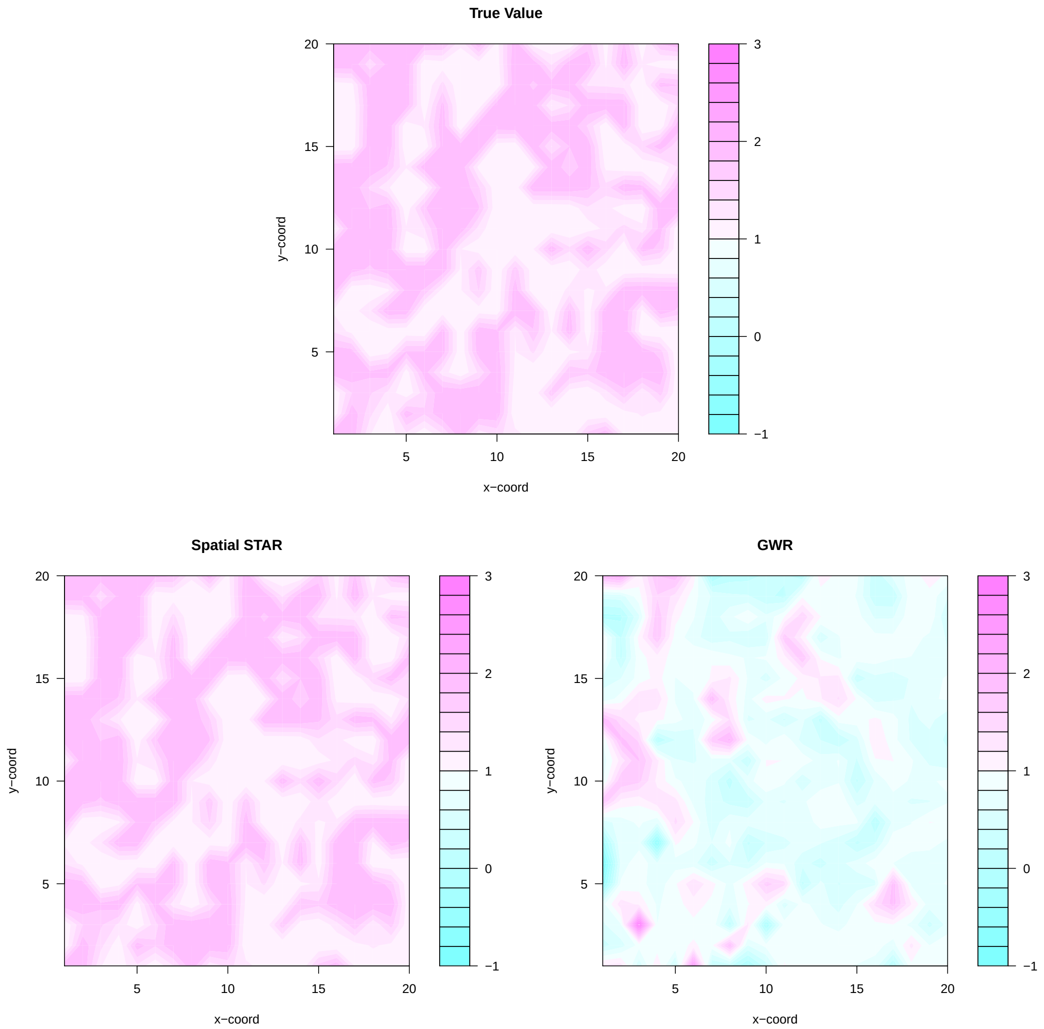

Figure 3: Contour Plot of β_0 ($\beta_0 + \delta_0 \circ G(\cdot)$ for the true value and spatial STAR)

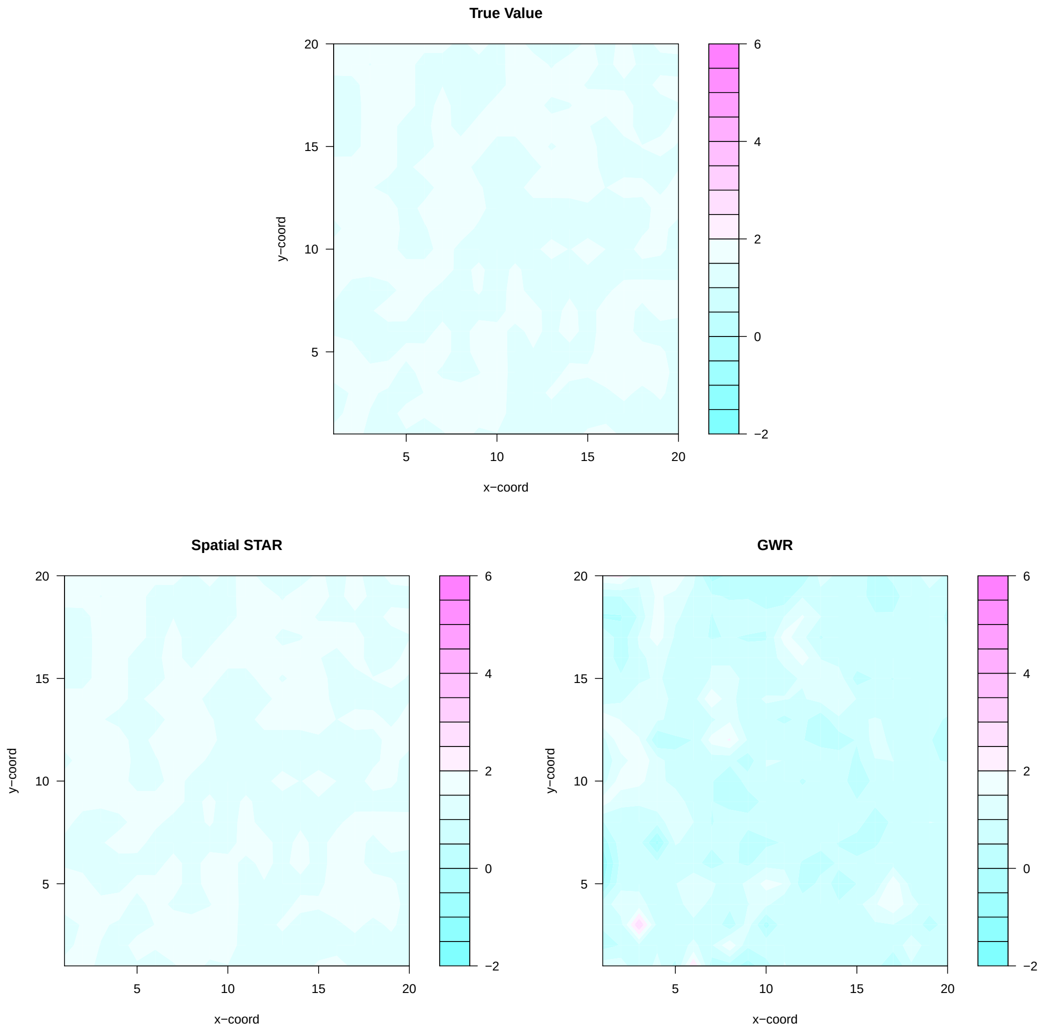

Figure 4: Contour Plot of β_1 ($\beta_1 + \delta_1 \circ G(\cdot)$ for the true value and spatial STAR)

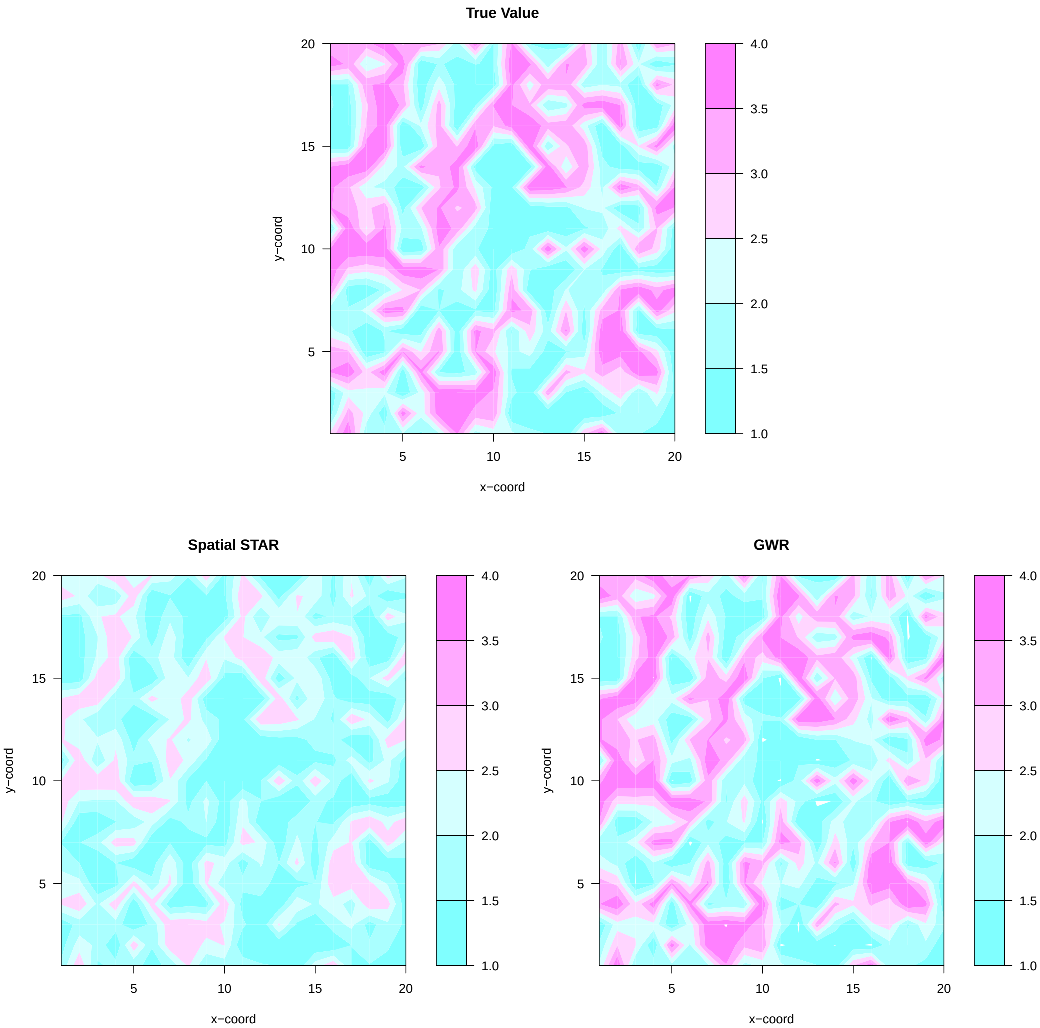

Figure 5: Contour Plot of $\widehat{y}$

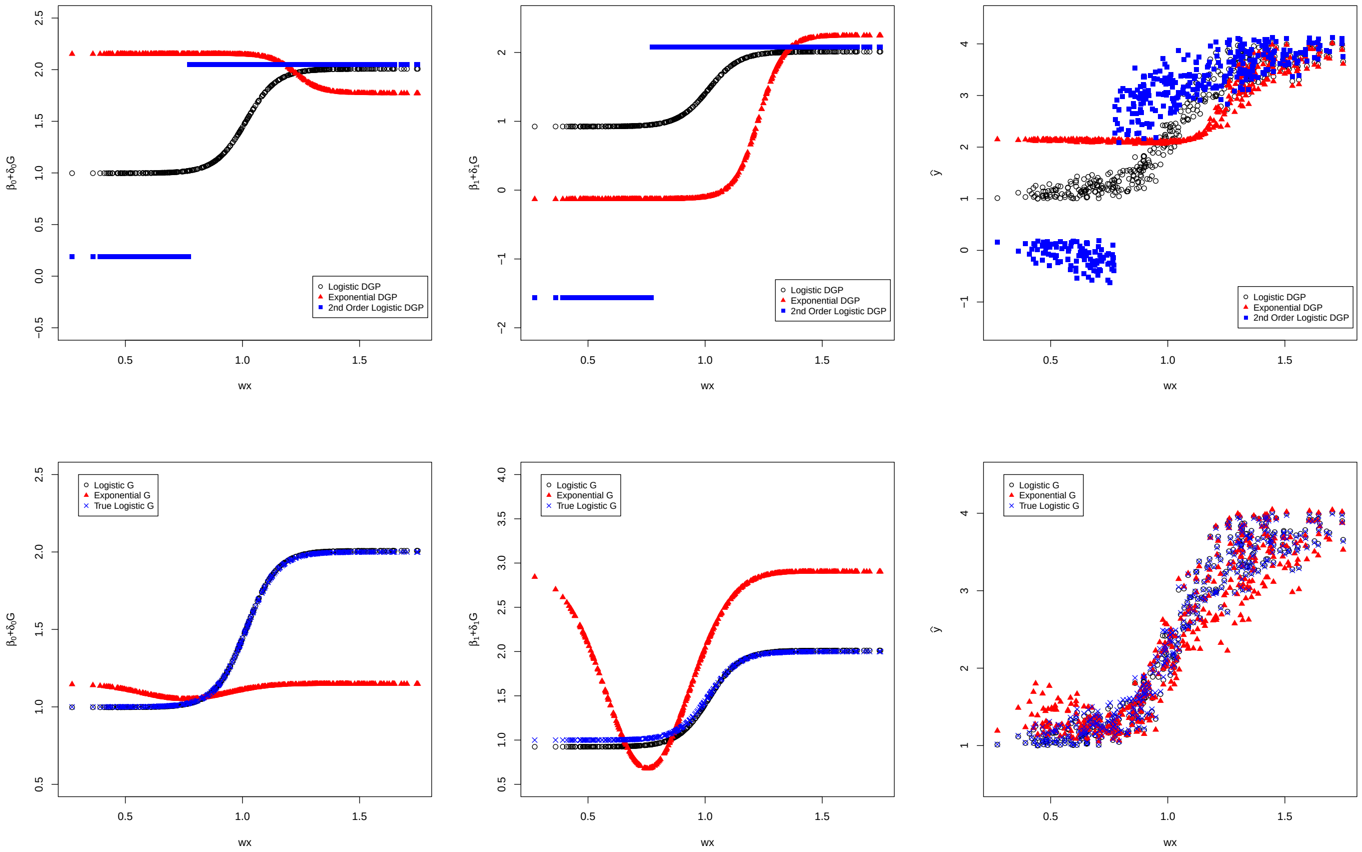

Figure 6: Scatter plots for Combinations of DGP and Estimation Function $G(\cdot)$: (Upper Panel) Different DGPs with Logistic $G(\cdot)$ Estimation Function (Lower Panel) Logistic DGP with Different $G(\cdot)$ Estimation Functions